1. INTRODUCTION

- Volume of data transmission for the users of personal communications and personal computer systems is growing unbounded, even to the extend of bringing the service providers to a complete stall.
- World-wide-web facilities have opened the eyes of millions of users to be more demanding on everything they interface with in their daily lives.
- Increasing demand for high quality digital telephony, digital TV, and multi-media communications over advanced networks prompted numerous studies in the area of data compression

This course will introduce both:

- Classical techniques in speech, image, and video compression
- Emerging algorithms and standards in the field
- Demonstrate through a number of applications by means of projects contributed by students.

Primary objective is to broaden general and working knowledge of engineers and scientist in the area of modern techniques on signal compression and in their implementations.

This course is designed for engineers and researchers involved in all fields of signal processing including:

- speech coding and transmission,
- image compression,
- video coding in multimedia applications,
- telecommunication systems and services,
- defense and
- manufacturing.

Communication System Models:

Classical Communication Systems Model:

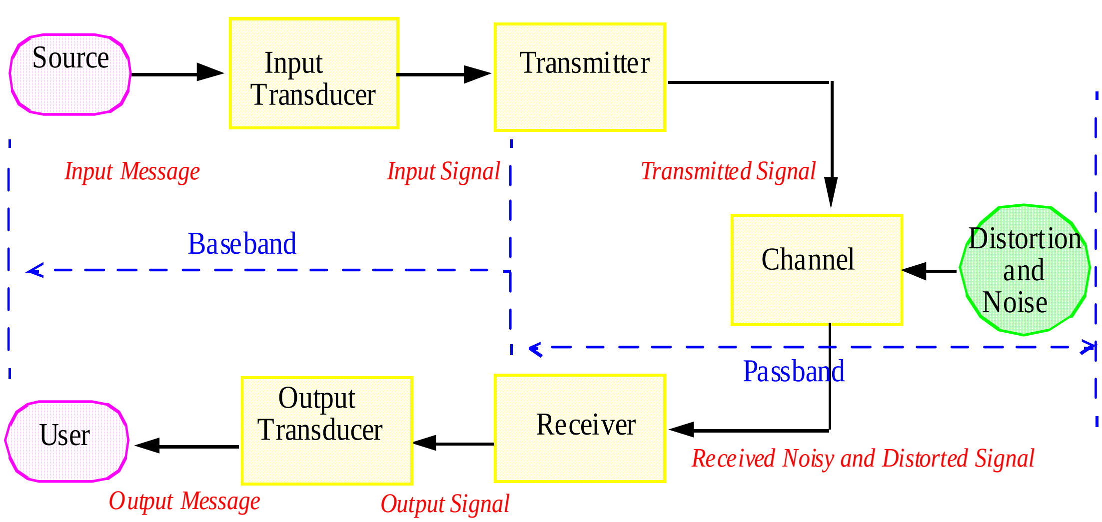


Figure 1.1. Basic Communication System Block Diagram

Shannon's Point-to-Point Digital Communication Systems Model

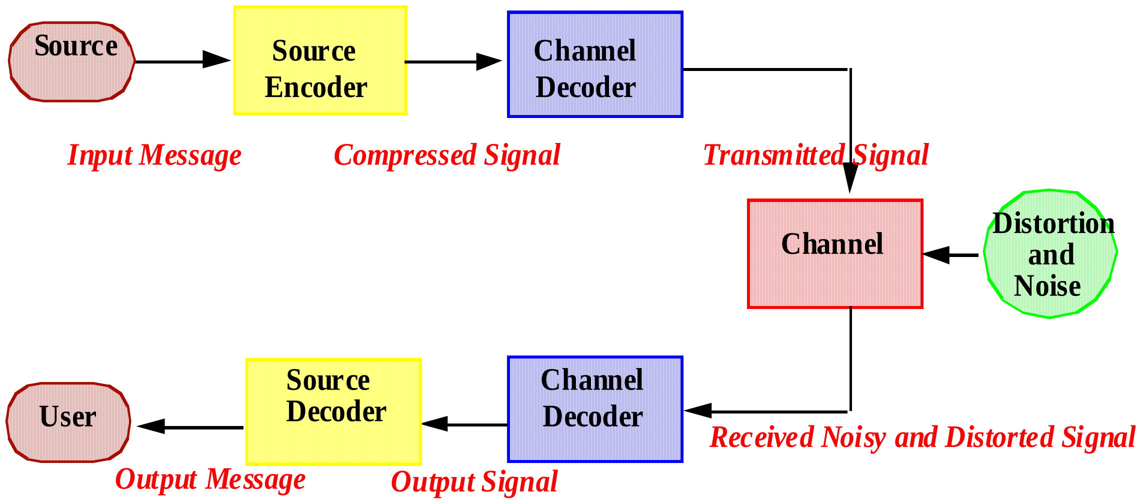


Figure 1.2. Shannon's Point-to-Point Digital Communication Model

Source:

Set of Symbols generated by a person or a system to be sent over a transmission medium to a user.

Examples

- Speech/audio
- Image/video
- Telemetry and other sensor data
- Computer dat

We normally model the source as a random process: $\{X_n\}$ or $\{X(t)\}$; $\{X_{n,m}\}$ or $\{X(t,s)\}$ depending upon discrete or continuous "time: $t, s \in \Re$" or "space: $n, m \in \aleph$" with a discrete or continuous alphabet: $X_n \in A \subset \Re^k$.

	Continuous Time/Space	**Discrete Time/Space**
Continuous Amplitude	Microphone, photos, x-ray, radar, earthquake sensor	Sampled waveform, features vectors (LPC, MFCC, PCA, ICA, eigenface
Discrete Amplitude	Neuron, digital thermometer	Computer files, internet data

Source Encoder

Messages from users are highly redundant. Compression of redundancy in a systematic manner is called source encoding.

Examples:

- CELP coding for speech/audio signals
- JPEG coding for still images
- Lempel-Ziv universal lossless coding for text compression

Channel Encoder

Coding for improved transmission over physical medium.

Examples:

- Run-length line coding
- Convolutional codes
- QAM (ADSL Modems), FSK (very-low-rate modems), DPSK (low-rate modems, QPSK (cable modems, CDMA codecs), and other codes for data transmission.

In many cases two are combined and called **"Encoder"** and described by: What designer gets to do to signal before sending it over the channel. It can include:

- Preprocessing,
- Sampling and A/D conversion,
- Signal decompositions,
- Modulation, and
- Compression.

Goal: Prepare signal for channel in a way decoder can recover good reproduction.

Channel

Physical medium for communication process. This portion of communication system is out of designer's control. It is often described in terms of a conditional probability distribution and a linear filtering operation. It could be: Deterministic or Random

Examples:

1. **On-line media:**
   - Null (transparent) channel
   - Air/deep space
   - Telephone lines; twisted-pair/coaxial cable
   - Ethernet
   - Fiber-optic line
2. **Off-line media:**
   - CD
   - Magnetic tape/Magnetic disk
   - Computer memory

Channel and Source Decoders

They attempt to perform inverse operations of the source encoder and the channel encoder, respectively. The combination is called as the **"Decoder"** and described by: What decoder gets to do to channel output in order to reconstruct or render a version of the signal for the user. It can include inverses or approximate inverses of encoder operations, or other stuff to enhance reproduction.

Distortion and Noise

When the continuous or analog signals are digitized and compressed there is always a cost associated with the process. In digitization of band-limited signals, we employ Nyquist Theorem to guarantee exact reconstruction. However, any other source compression is realized at a cost of varying degree of imperfect representation. This is called distortion and it is NOT recoverable. In addition, signals in the communication link are faced with number of ills. They are loosely called noise.

The presence of noise on a signal changes it shape and characteristics and it limits the ability of the intended receiver to make correct symbol decisions, and thereby the effects the rate of reliable communication.

Examples:

- Additive Gaussian White Noise
- Device noise
- Atmospheric noise in the microwave channels
- Intersymbol interference in data communication systems
- Interspeaker interference in voice communications
- Near-end and Far-end crosstalk
- Echoes in Link and chamber
- Friendly and unfriendly jammers, etc.

User or Application

The intended user of the input information-bearing message, usually a replica of the original input message.

The messages coming to user may not need an identical replica of the sender's information symbols, instead, it can end up as an action item (application).

Example:

- The access control of a safe room by the voice print of intended user.

2. INTRODUCTION TO IMAGE PROCESSING

Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers. The goal of this manipulation can be divided into three categories:

*** Image Processing** ***image in -> image out***

*** Image Analysis** ***image in -> measurements out***

*** Image Understanding** ***image in -> high-level description out***

We will focus on the fundamental concepts of *image processing*. Further, we will restrict ourselves to two-dimensional (2D) image processing although most of the concepts and techniques that are to be described can be extended easily to three or more dimensions.

We begin with certain basic definitions:

Image: image is considered to be a function of two real variables, for example, $a(x,y)$ with a as the amplitude (e.g. brightness) of the image at the *real* coordinate position (x,y). The amplitudes of a given image will almost always be either real numbers or integer numbers.

ROI: An image may be considered to contain sub-images sometimes referred to as *regions-of-interest*, *ROIs*, or simply *regions*.

This concept reflects the fact that images frequently contain collections of objects each of which can be the basis for a region. Thus one part of an image (region) might be processed to suppress motion blur while another part might be processed to improve color rendition.

Quantization: The latter is usually a result of a quantization process that converts a continuous range (say, between 0 and 100%) to a discrete number of levels. In certain image-forming processes, however, the signal may involve photon counting which implies that the amplitude would be inherently quantized. In other image forming procedures, such as magnetic resonance imaging, the direct physical measurement yields a complex number in the form of a real magnitude and a real phase. For the remainder of this course we will consider amplitudes as reals or integers unless otherwise indicated.

Digital Image: A digital image $a[m,n]$ described in a 2D discrete space is derived from an analog image $a(x,y)$ in a 2D continuous space through a *sampling* process that is frequently referred to as digitization. For now we will look at some basic definitions associated with the digital image. The effect of digitization is shown in Figure 1.3. The 2D continuous image $a(x,y)$ is divided into N *rows* and M *columns*. The intersection of a row and a column is termed a *pixel*. The value assigned to the integer coordinates $[m,n]$ with $\{m=0,1,2,...,M-1\}$ and $\{n=0,1,2,...,N-1\}$ is $a[m,n]$. In fact, in most cases $a(x,y)$--which we might consider to be the physical signal that impinges on the face of a 2D sensor--is actually a function of many variables including depth (z), color (λ), and time (t). Unless otherwise stated, we will consider the case of 2D, monochromatic, static images in this chapter.

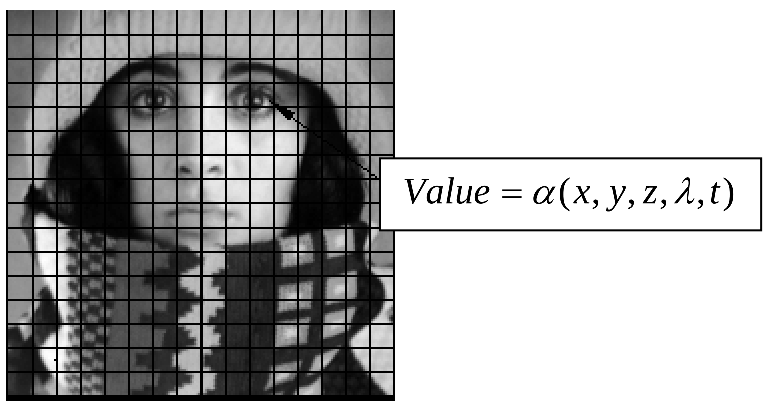
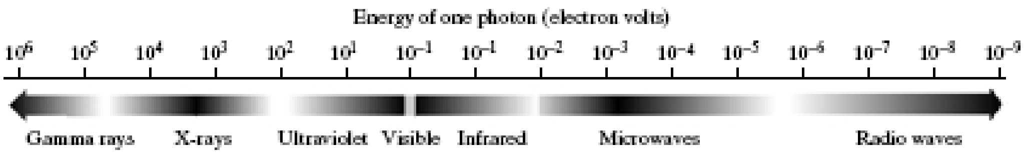


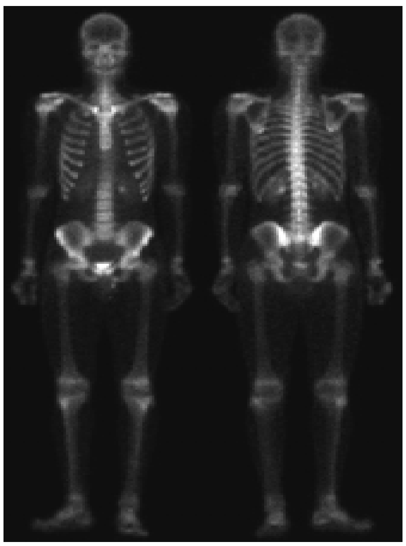
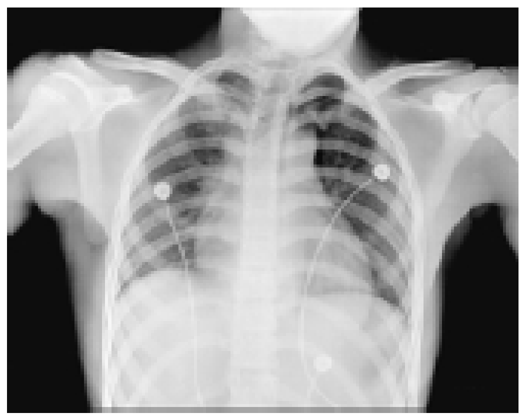
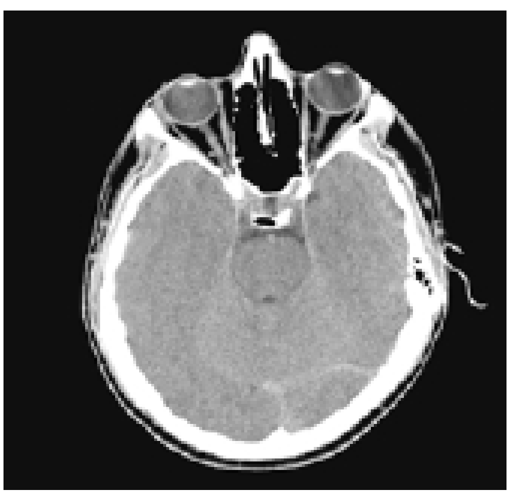
Figure 1.3: Digitization of a continuous image. The pixel at coordinates [m=10, n=3] has the integer brightness value 110.

The image shown in Figure 1.3 has been divided into N = 16 rows and M = 16 columns. The value assigned to every pixel is the average brightness in the pixel rounded to the nearest integer value. The process of representing the amplitude of the 2D signal at a given coordinate as an integer value with L different gray levels is usually referred to as amplitude quantization or simply *quantization*.

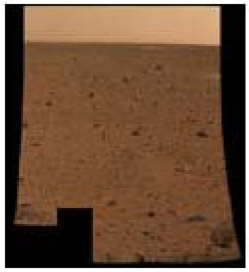
Images are formed from the **radiation of electromagnetic (EM) spectrum** both in the visual and non-visual bands (Gamma rays through radio waves) as shown below (GW: Figure 1.5.)



Examples of Different Imagery: Gamma and X-Ray Imagery


Some infrared images from Mars Rover "Spirit", which are used for aligning the landing trajectory of the vehicle (Images are downloaded form the NASA/JPL Website.)



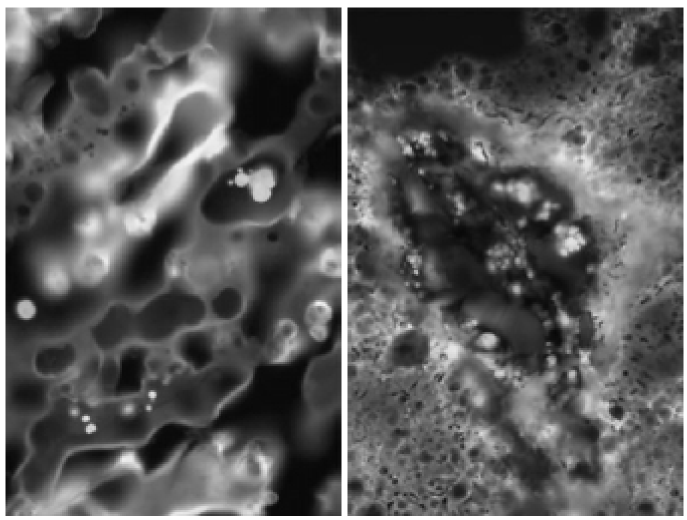
Color mosaic image
Of Mars surface

taken by
Mars Global Surveyor

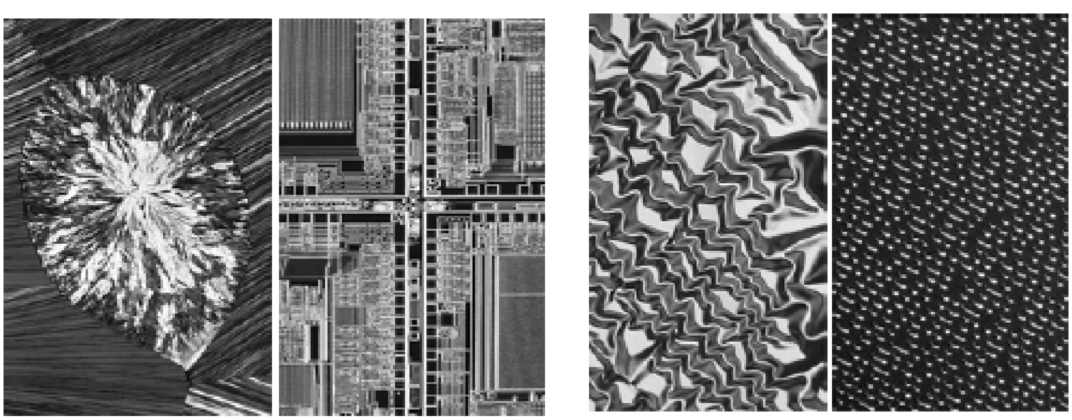
taken by
Rover's descent imaging
motion estimation
(DIME) system

Ultraviolet Imaging used in many industrial inspections, microscopy and biological imaging:

Normal and smut infected corn



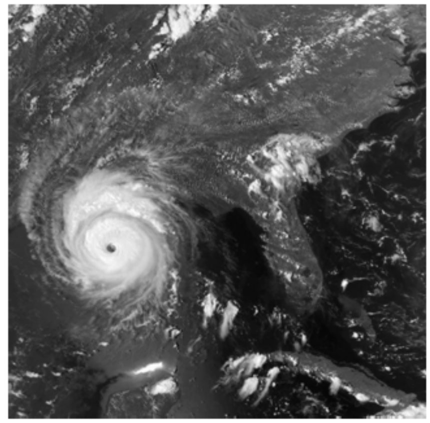
Visible and Infrared Imagery



From left : Cholestrol(40x); Microprocessor (60x); Audio CD surface(1750x); Organic superconductor(450x).

Thematic Imagery from Landsat Satellite Imaging. There are seven thematic bands used by NASA ranging from 0.45 μm (visible blue) to 12.25 μm thermal infrared, used in many applications.

Band No.	Name	Wavelength (μm)	Characteristics and Uses
1	Visible blue	0.45–0.52	Maximum water penetration
2	Visible green	0.52–0.60	Good for measuring plant vigor
3	Visible red	0.63–0.69	Vegetation discrimination
4	Near infrared	0.76–0.90	Biomass and shoreline mapping
5	Middle infrared	1.55–1.75	Moisture content of soil and vegetation
6	Thermal infrared	10.4–12.5	Soil moisture; thermal mapping
7	Middle infrared	2.08–2.35	Mineral mapping



Seven Thematic Bands and Multi-band imagery of the eye of Hurricane Andrew.

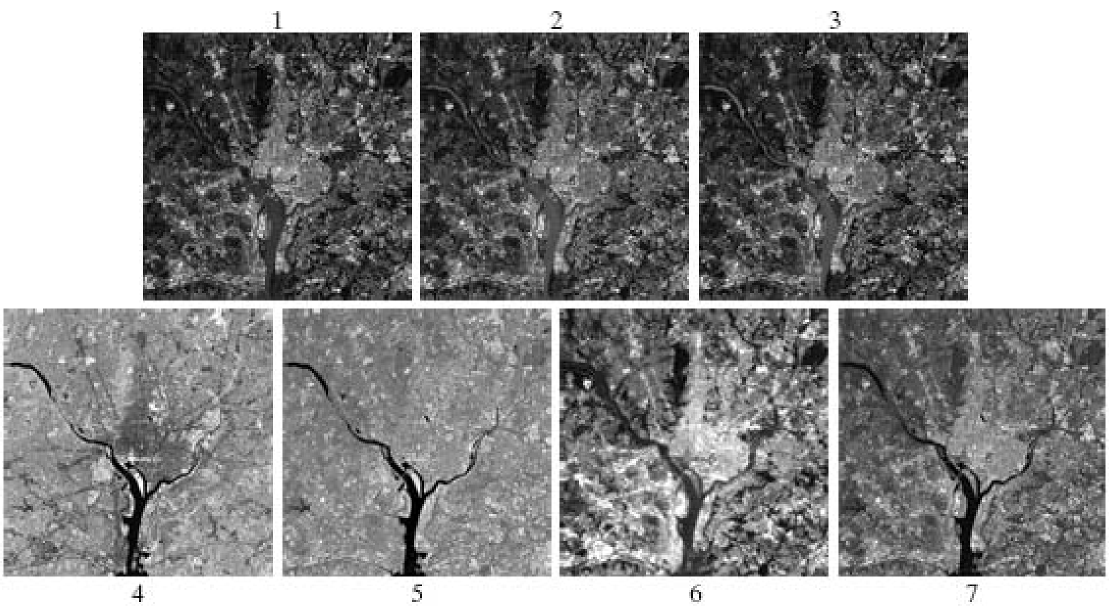
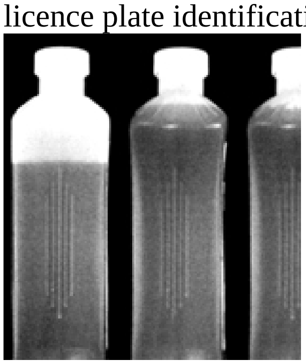
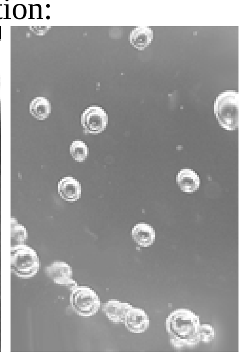
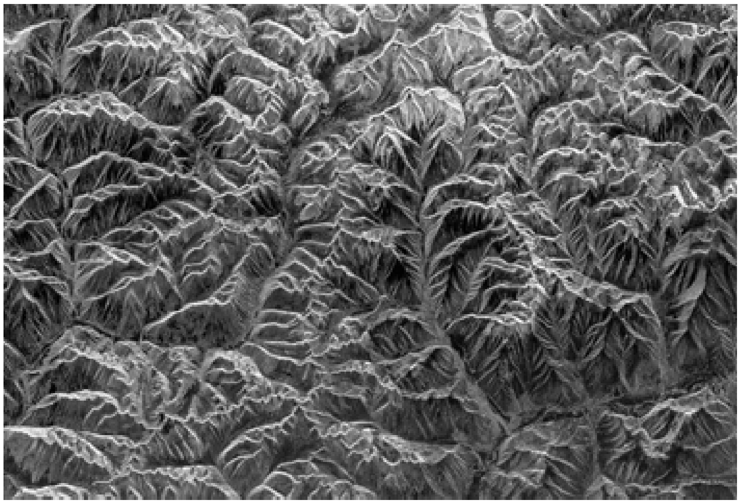


FIGURE 1.10 LANDSAT satellite images of the Washington, D.C. area. The numbers refer to the thematic bands in Table 1.1. (Images courtesy of NASA.)

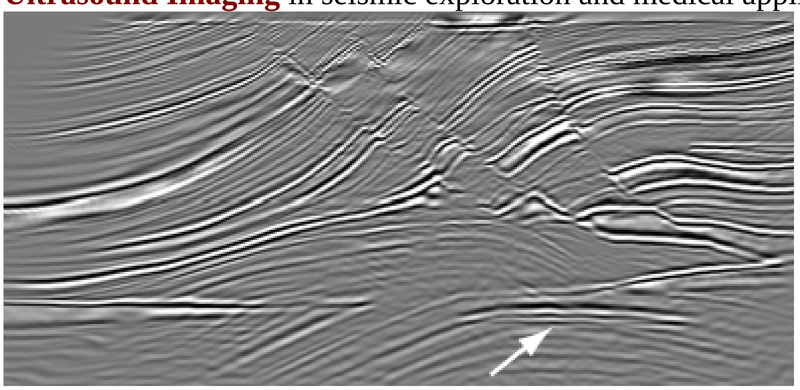
Industrial and Law Enforcement Applications: Bottle and air pocket inspection, fingerprint and licence plate identification:

Imaging in the radio band: Spaceborne radar image of Southeast Tibet NS MRI Image of a knee.



Ultrasound Imaging in seismic exploration and medical applications:



Common Values encountered in Digital Image Processing

There are standard values for the various parameters encountered in digital image processing. These values can be caused by video standards, by algorithmic requirements, or by the desire to keep digital circuitry simple. Table 1 gives some commonly encountered values.

Parameter	*Symbol*	*Typical values*
Rows	N	256,512,525,625,1024,1035
Columns	M	256,512,768,1024,1320
Gray Levels	L	2,64,256,1024,4096,16384

Table 1.1: Common values of digital image parameters

Quite frequently we see cases of $M=N=2^K$ where $\{K = 8,9,10\}$. This can be motivated by digital circuitry or by the use of certain algorithms such as the (fast) Fourier transform (see Section 3.3).

The number of distinct gray levels is usually a power of 2, that is, $L=2^B$ where B is the number of bits in the binary representation of the brightness levels. When $B>1$ we speak of a *gray-level image*; when $B=1$ we speak of a *binary image*. In a binary image there are just two gray levels which can be referred to, for example, as "black" and "white" or "0" and "1".

Characteristics of Image Operations: There is a variety of ways to classify and characterize image operations. The reason for doing so is to understand what type of results we might expect to achieve with a given type of operation or what might be the computational burden associated with a given operation.

Types of operations: The types of operations that can be applied to digital images to transform an input image a[m,n] into an output image b[m,n] (or another representation) can be classified into three categories as shown in Table 2, which is shown graphically in Figure 2.

Operation	Characterization	Generic Complexity/Pixel
* *Point*	- the output value at a specific coordinate is dependent only on the input value at that same coordinate.	*constant*
* *Local*	- the output value at a specific coordinate is dependent on the input values in the *neighborhood* of that same coordinate.	P^2
* *Global*	- the output value at a specific coordinate is dependent on all the values in the input image.	N^2

Table 1.2: Types of image operations. Image size = N x N; neighborhood size = P x P. Note that the complexity is specified in operations *per pixel*.

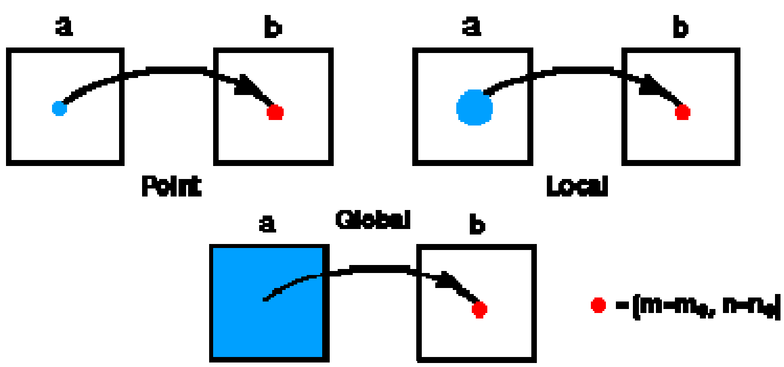


Figure 1.4: Illustration of various types of image operations